

## Report

# Voice-Sensitive Regions in the Dog and Human Brain Are Revealed by Comparative fMRI

Attila Andics,<sup>1,\*</sup> Márta Gácsi,<sup>1</sup> Tamás Faragó,<sup>1</sup> Anna Kis,<sup>2,3</sup> and Ádám Miklósi<sup>1,2</sup>

<sup>1</sup>MTA-ELTE Comparative Ethology Research Group, Pázmány Péter sétány 1/C, 1117 Budapest, Hungary

<sup>2</sup>Department of Ethology, Eötvös Loránd University, Pázmány Péter sétány 1/C, 1117 Budapest, Hungary

<sup>3</sup>Research Centre for Natural Sciences, Institute of Cognitive Neuroscience and Psychology, Hungarian Academy of Sciences, Magyar tudósok körútja 2, 1117 Budapest, Hungary

## Summary

During the approximately 18–32 thousand years of domestication [1], dogs and humans have shared a similar social environment [2]. Dog and human vocalizations are thus familiar and relevant to both species [3], although they belong to evolutionarily distant taxa, as their lineages split approximately 90–100 million years ago [4]. In this first comparative neuroimaging study of a nonprimate and a primate species, we made use of this special combination of shared environment and evolutionary distance. We presented dogs and humans with the same set of vocal and nonvocal stimuli to search for functionally analogous voice-sensitive cortical regions. We demonstrate that voice areas exist in dogs and that they show a similar pattern to anterior temporal voice areas in humans. Our findings also reveal that sensitivity to vocal emotional valence cues engages similarly located nonprimary auditory regions in dogs and humans. Although parallel evolution cannot be excluded, our findings suggest that voice areas may have a more ancient evolutionary origin than previously known.

## Results and Discussion

An important social function of the auditory system is to process the vocalizer's identity and emotional state. Nonprimary auditory brain regions preferring conspecific vocalizations were found in both humans [5, 6] and nonhuman primates, suggesting that “voice areas” evolved at least 30 million years ago [7–10]. In humans, auditory regions sensitive to vocal emotional cues have also been identified [11–14]. Research has also indicated that vocal emotional valence is conveyed via similar acoustic rules across species [15], including human [16] and nonhuman [17] animals.

Behavioral field research has revealed that the efficient processing of conspecificity and emotional information in vocalizations is important in both primate [18–20] and nonprimate [21–24] species. Indeed, both the acoustic recognition of conspecifics and their emotional state are fundamental for making decisions in behavior contexts like mate choice, territory disputes, or hierarchy-related challenges [25]. Nevertheless, little is known about the underlying neural mechanisms of vocalization processing in nonprimates.

To reveal possible functional analogies between human and nonprimate auditory brain regions, this study describes a comparative investigation of dogs and humans. We investigated (1) whether in dogs, similarly to humans, certain auditory regions (voice areas) would respond stronger to conspecific vocalizations than to either heterospecific vocalizations or nonvocal sounds and (2) whether dogs are similar to humans in the cortical processing of emotional cues in vocal signals.

To address these questions, we used a noninvasive functional magnetic resonance imaging (fMRI) procedure with awake dogs ( $n = 11$ ) and humans ( $n = 22$ ). We built on our group's first small sample-sized attempts of awake dog fMRI [26], providing a procedure different from others' [27]. All participants were unrestrained and instructed to lay motionless in an fMRI scanner for three 6 min runs (see [Figure 1](#) and [Figure S1](#) available online). Dogs and humans listened to an identical set of stimuli, which included three sound types: human vocalizations, dog vocalizations, nonvocal environmental sounds, and a silent baseline. Vocal stimuli ranged parametrically in emotional valence from highly negative to highly positive, as rated by an independent set of human listeners [16]. Neural sensitivity to conspecificity and emotional valence (and related acoustic cues) was evaluated similarly for the two species using random-effects group analyses (see the [Supplemental Experimental Procedures](#)).

Auditory regions were defined functionally, using the all sounds versus silence contrast ([Figures 2A](#) and [2B](#)). Similar sound-sensitive brain regions were identified in dogs and humans, including regions of the auditory cortex and subcortical regions ([Table S1](#)). Consistent with lesion studies on dog auditory processing regions [28, 29], cortical sound sensitivity in dogs was localized in perisylvian regions, including the Sylvian gyri (SG) along the Sylvian fissure (SF) and the ectosylvian gyri (ESG) along the ectosylvian sulcus (ESS), and extending dorsally to the suprasylvian sulcus (SSS). Human cortical auditory activity was found along the superior temporal sulcus (STS) and in the inferior frontal cortex (IFC). In both species, auditory activity extended ventrally toward the temporal pole (TP), i.e., the most basal part of the caudal SG in dogs, and the anterior tip of the temporal lobe in humans. A subcortical sound-sensitive region, with a peak in the first-order auditory thalamus, the medial geniculate body (MGB, [30, 31]), including the caudal colliculus and extending toward the cerebellum, was also identified in both species. The search space of all following analyses was defined by these functionally localized auditory regions: their total size was 12 cm<sup>3</sup> (1,441 voxels) for dogs and 95 cm<sup>3</sup> (11,849 voxels) for humans.

A first qualitative comparison of parameter estimates for each sound type indicates an important difference between dog and human auditory regions. Dogs have subregions in which parameter estimates were maximal for dog vocalizations (39% of all auditory voxels), but also subregions with maximal response to human vocal (13%) or nonvocal (48%) sounds. In contrast, almost all human auditory regions were maximal for human vocalizations (87%). Maximal response for dog vocalizations (10%) was found in the subcortical MGB, and almost no subregions were found where the response was maximal for nonvocal sounds (3%) ([Figure 2C](#)).

\*Correspondence: [attila.andics@gmail.com](mailto:attila.andics@gmail.com)



Figure 1. Steps of Positioning a Dog in the fMRI Scanner

(A) Dog lying on scanner bed, being rewarded with food and socially by the owner.

(B) As part of the model-rival training procedure, another dog is observing as the tested dog is praised while receiving earphones from an experimenter.

(C) When the upper element of the coil is fixed with stripes on the top of the dog's head, the scanner bed is moved to the scanning position. See also the [Supplemental Experimental Procedures](#), [Figure S1](#), and [Movie S1](#).

To identify voice areas, i.e., auditory regions responding preferentially to conspecific vocalizations compared to either heterospecific or nonvocal sounds, we compared brain responses to each sound type in random-effects conjunction analyses ([Figure 3](#), [Table S2](#), and the [Supplemental Experimental Procedures](#)). In dogs, we identified a ventral auditory region (cSG), close to the TP bilaterally, and a left dorsal auditory region (mESG) that responded stronger to dog than to either human or nonvocal sounds. None of these regions responded stronger to human than to nonvocal sounds. In humans, regions along the bilateral STS, including posterior (pSTS), mid (mSTS), and anterior (aSTS) STS, extending to the TP, and also the right IFC, were more sensitive to human than to either dog or nonvocal sounds. There was also a difference in the relation of the nonpreferred dog and nonvocal sounds across the temporal subregions in humans. While in pSTS and mSTS the response to dog sounds was between that to human and nonvocal sounds, aSTS and TP regions showed no preference for dog compared to nonvocal sounds: repeated-measures ANOVAs with factors anteriority (pSTS, mSTS, [right] aSTS, TP) and sound type (dog, nonvocal) showed a significant interaction of the two factors for each hemisphere [left:  $F(2,42) = 45.386$ ,  $p < 0.001$ ; right:  $F(3,63) = 42.491$ ,  $p < 0.001$ ]. We also looked for regions responding stronger to heterospecific vocalizations than to other sound types. In dogs, no regions showed stronger responses to human than either dog or nonvocal sounds. In humans, only the subcortical MGB, but no temporal regions, showed greater sensitivity to dog than either human or nonvocal sounds.

These findings provide the first evidence for the existence of voice areas in dogs, or in any nonprimate brain. The only bilateral conspecific-prefering region in dogs is near the TP, i.e., the ventral part of the caudal SG, extending to the SF. In humans, the anterior voice areas (i.e., aSTS and TP) are special in that, similarly to dog voice areas, they respond most strongly to conspecific sounds, but do not respond stronger

to heterospecific than to nonvocal sounds. These similarly located (i.e., near the TP) auditory cortex regions thus appear to be functionally analogous in dog and human brains. Anterior temporal and TP regions have been implied in conspecific vocalization processing in both nonhuman primates [7–10, 32, 33] and humans [6, 34, 35]. More specifically, these regions have been implied in voice identity processing, a key function of voice areas [6, 35, 36]. While claims about exact

anatomical correspondences and therefore about homologies across dog and human brain regions are difficult to make and are beyond the scope of this paper, a plausible interpretation of our findings is that conspecific preference in these auditory regions is an evolutionarily ancient function across mammalian orders, although convergent evolution [37] is an alternative. At the very least, our results show that, similarly to primates, conspecific vocalizations have a special status in the dog brain.

In humans, consistent with previous reports [5, 6, 35], the temporal voice areas involved not only anterior regions, but also the mid and posterior STS. Here we show that pSTS and mSTS, unlike aSTS and TP, prefer heterospecific (dog) sounds to nonvocal sounds (cf. [38]). This suggests that pSTS and mSTS are not strictly conspecific specific, but rather tuned to familiar, relevant vocal sounds in general, an interpretation possibly also supported by a report finding no preference for conspecific compared to human vocalizations in the macaque mSTS [32]. Replicating earlier findings, we also found conspecific preference in the human right IFC [39, 40], another region implied in voice identity processing [40, 41]. Finally, we found that, in humans, the subcortical MGB, previously implied in processing rapidly varying spectrotemporal features of human vocal sounds [42], responded stronger to dog sounds than other sound types.

We also tested, in a series of parametric modulation analyses, whether the emotional valence of vocalizations is reflected in brain responses and how such responses are modulated by acoustical cues. Vocal stimuli were blocked by valence scores (ranging from highly negative to highly positive). The affective context valence of the dog vocalization recordings was found to covary with human emotional valence ratings of these sounds, suggesting that human ratings represent a fairly good evaluation of the animal's affective state ([16] and the [Supplemental Experimental Procedures](#)). Block-averaged valence scores were then used as parametric

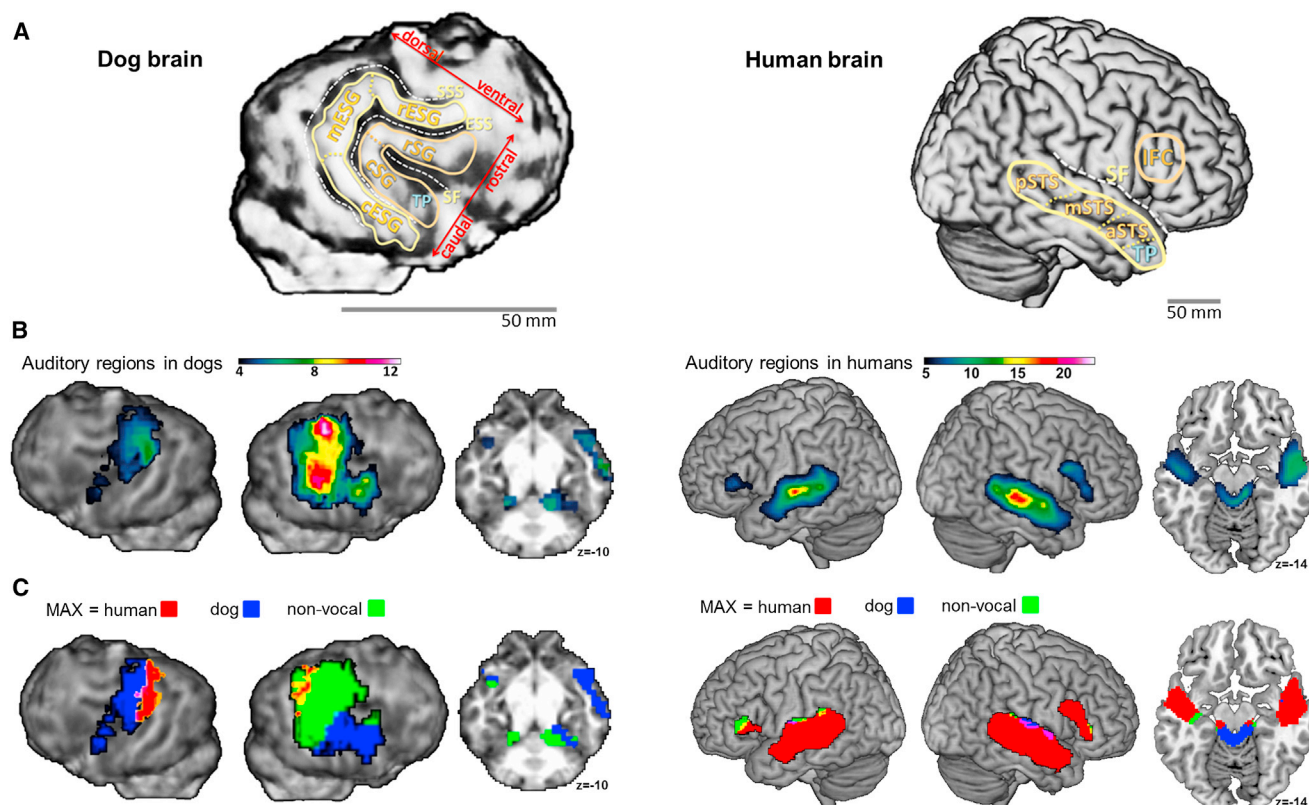


Figure 2. Auditory Regions in Dogs and Humans

(A) Schematic representations of sound-sensitive perisylvian regions in dogs and humans, superimposed on rendered brains. Dog abbreviations are as follows: c, caudal; m, middle; r, rostral; ESG, ectosylvian gyrus; ESS, ectosylvian sulcus; SF, Sylvian fissure; SG, Sylvian gyrus; SSS, suprasylvian gyrus; and TP, temporal pole. Human abbreviations are as follows: a, anterior; m, mid; p, posterior; IFC, inferior frontal cortex; SF, Sylvian fissure; STS, superior temporal sulcus; and TP, temporal pole.

(B) Auditory regions as determined by the all sounds versus silence contrast in dogs and humans, thresholded at  $p < 0.001$ , FEW corrected at the cluster level, using the uncorrected voxel threshold  $p < 0.001$  for dogs (in a whole-volume search space of  $90 \text{ cm}^3$ ) and  $p < 0.00001$  for humans (in a whole-volume search space of  $1,277 \text{ cm}^3$ ). Color heatmaps indicate  $t$  values, superimposed on rendered brains and selected axial slices.

(C) The same auditory maps as in (B). The color code refers to the sound type that elicited the maximal response in each voxel.

See also Table S1 and Audio S1.

modulators to test whether auditory brain activity covaries with emotional content. Specifically, we tested whether emotional vocalizations that are perceived as more positive (in a parametric manner) elicit greater (or smaller) neural responses (for details, see the [Supplemental Experimental Procedures](#)).

Emotional valence-sensitive regions were identified in both dogs and humans (Figure 3). These regions all responded stronger to more positive vocalizations—we found no regions responding stronger to more negative vocalizations. In dogs, we found that an auditory region in the right cESG, close to the primary auditory cortex, was sensitive to emotional valence, for both dog and human vocalizations (Table S3). No emotional modulation effect was found in the corresponding cESG region of the left hemisphere, with a significant difference across hemispheres [ $T(10) = 2.234$ ,  $p < 0.05$ , see the [Supplemental Experimental Procedures](#)]. We found a similar, but weaker, modulatory effect of emotional valence in a bilateral rostral SG region for human, but not dog, vocalizations. In humans, an analogous effect was found: neural activity in the auditory cortex, with a maximum in the mSTS, increased with the perceived emotional positivity of vocalizations. This emotional modulation effect was present bilaterally

for both dog and human vocal stimuli, with only a tendency for a right-hemisphere bias [ $T(21) = 1.879$ ,  $p = 0.074$ ]. The human mSTS has been implied in the extraction of social or affective salient signals from conspecific vocalizations [11]. Our results show that the same mechanism may be used to extract affective information from heterospecific vocalizations. Furthermore, dogs appear to use a similar mechanism, localized in the cESG, for extracting vocal emotional information from either conspecifics or humans.

Additionally, we tested how auditory regions in each species are modulated by acoustic parameters relevant for emotional processing. In a related paper [16], we already established that perceived emotional valence and intensity of these vocal stimuli covary with a basic temporal cue (call length) and a basic spectral cue (fundamental frequency, F0), respectively. Specifically, emotional valence increases with decreasing call length, while emotional intensity increases with increasing F0. Here we found that auditory regions are parametrically modulated (1) by call length in dogs (activity decreased with increasing call length in the bilateral mESG and also in the subcortical MGB) and (2) by F0 in both species (activity decreased with increasing F0, in the right mESG and in a left rSG region in dogs, and in the m/pSTS in humans)



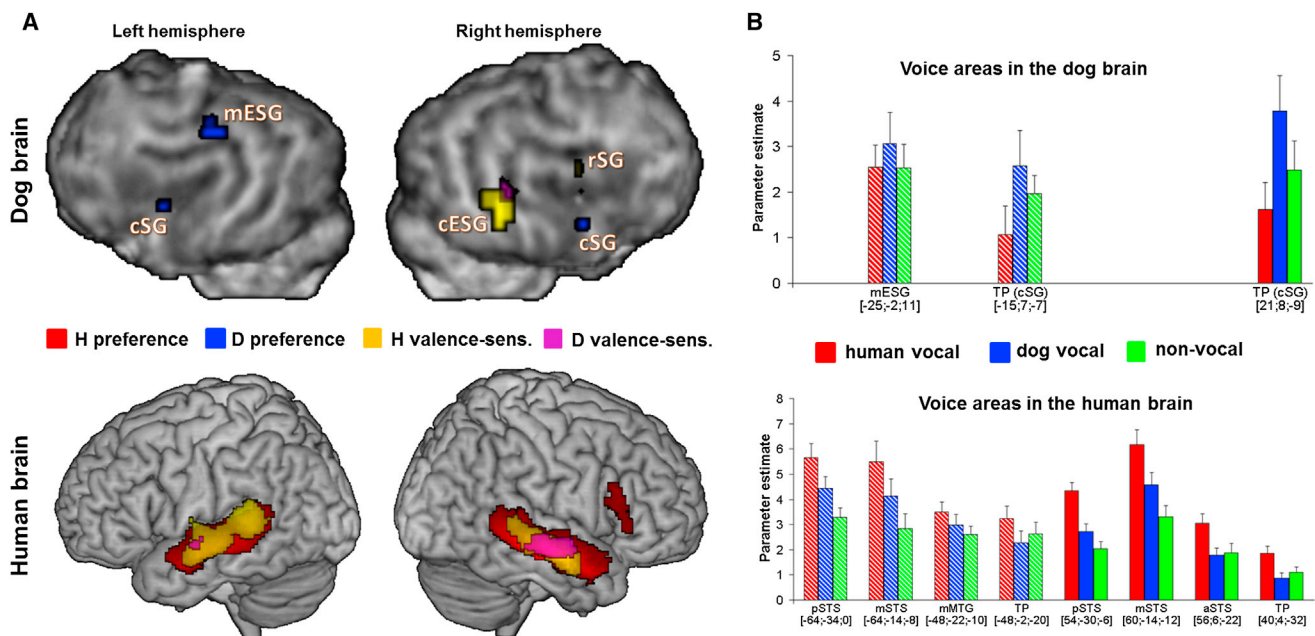


Figure 3. Species Preference and Emotional Valence Sensitivity for Vocalizations in Dogs and Humans

(A) Activity maps, superimposed on rendered brains, are thresholded at  $p < 0.005$  for dogs and at  $p < 0.0005$  (in clusters of at least 10 voxels) for humans. Regions with human preference (a conjunction of human vocal > dog vocal and human vocal > nonvocal; red), dog preference (a conjunction of dog vocal > human vocal and dog vocal > nonvocal; blue), human valence sensitivity (a positive parametric effect of valence of human vocal sounds; yellow), and dog valence sensitivity (a positive parametric effect of valence of dog vocal sounds; purple) are shown.

(B) Parameter estimates for voice area peaks. Bars represent beta weights for each sound type. Striped bars, left-hemisphere peaks; filled bars, right-hemisphere peaks. Error bars indicate the SEM.

See also Tables S2 and S3.

(Table S3). In both dogs and humans, the right auditory cortex peaks for valence sensitivity and each acoustic parametric effect were within a 16 mm distance and close to primary auditory regions.

These findings suggest that acoustical cues related to vocal emotional valence are processed similarly in the dog and human auditory cortex. The involvement of a relatively early stage in the processing hierarchy in both species indicates that valence sensitivity at least partly reflects sensitivity in both species to acoustic parameters that convey emotions through voice. This is consistent with earlier human findings that imply emotional voice-sensitive regions in the mSTS and pSTS in processing both temporal [43] and spectral [12, 13, 31, 44] cues.

These discoveries suggest that the extraction of emotional information from voices is an important stage of the vocal emotion processing hierarchy and is supported by functionally analogous auditory brain regions near the primary auditory cortex in dogs and humans. These results expand earlier findings that dogs react similarly to some emotional state changes of other dogs and humans [45] and that humans recruit similar brain regions to process human and animal affective vocalizations [14]. These results also demonstrate that right-hemisphere dominance in vocal emotion processing, while debated in humans [46], is present in dogs, suggesting that behavioral lateralization effects in dog auditory processing [47] may be caused primarily by modulation of right-hemisphere activity.

This fMRI study compared, for the first time, two phylogenetically distant mammalian species under almost identical experimental conditions. Our results suggest common functions in dog and human voice processing. We presented

evidence that voice areas preferring conspecific vocalizations exist not only in primates, but also in dogs, and that, as in nonhuman primates [7–10, 32, 33] and humans [6, 34, 35], the dog voice areas involve bilateral TP regions. This evidence opens up the possibility that voice areas may have a longer evolutionary history than previously proposed [10], dating back to the common ancestor of dogs and humans some 100 million years ago [4], although convergent evolution cannot be excluded [37]. We also identified similarly located (i.e., near the primary auditory cortex) regions sensitive to emotional valence in vocalizations in both species and showed that this valence sensitivity involves keeping track of basic acoustic cues that mediate vocal emotions. This may be the first direct evidence suggesting that voice processing in mammalian listeners corresponds to the structural-functional organization of vocalizations [15] and forms the basis for using key acoustic features for cross-specific call recognition.

#### Supplemental Information

Supplemental Information includes Supplemental Experimental Procedures, one figure, three tables, one movie, and one audio file and can be found with this article online at <http://dx.doi.org/10.1016/j.cub.2014.01.058>.

#### Acknowledgments

The study was supported by the Hungarian Academy of Sciences (F01/031), the Hungarian Scientific Research Fund (OTKA K100695), and the ESF Research Networking Programme “CompCog”: The Evolution of Social Cognition (<http://www.compccog.org>; 06-RNP-020). We thank Gábor Rudas and the MR Research Centre of the Semmelweis University Budapest; József Topál, Viktor Gál, Zita Polgár, and four anonymous reviewers for comments on the manuscript; Ferenc Szalay for advice on dog brain

anatomy; and Katalin Czigány, Borbála Ferenczy, Bernadett Miklósi, and all owners for training the dogs.

Received: October 17, 2013

Revised: December 23, 2013

Accepted: January 28, 2014

Published: February 20, 2014

## References

- Thalmann, O., Shapiro, B., Cui, P., Schuenemann, V.J., Sawyer, S.K., Greenfield, D.L., Germonpré, M.B., Sablin, M.V., López-Giráldez, F., Domingo-Roura, X., et al. (2013). Complete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. *Science* 342, 871–874.
- Miklósi, A., and Topál, J. (2013). What does it take to become 'best friends'? Evolutionary changes in canine social competence. *Trends Cogn. Sci.* 17, 287–294.
- Pongrácz, P., Molnár, C., and Miklósi, Á. (2010). Barking in family dogs: an ethological approach. *Vet. J.* 183, 141–147.
- Springer, M.S., Murphy, W.J., Eizirik, E., and O'Brien, S.J. (2003). Placental mammal diversification and the Cretaceous-Tertiary boundary. *Proc. Natl. Acad. Sci. USA* 100, 1056–1061.
- Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., and Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature* 403, 309–312.
- Andics, A., McQueen, J.M., Petersson, K.M., Gál, V., Rudas, G., and Vidnyánszky, Z. (2010). Neural mechanisms for voice recognition. *Neuroimage* 52, 1528–1540.
- Petkov, C.I., Kayser, C., Steudel, T., Whittingstall, K., Augath, M., and Logothetis, N.K. (2008). A voice region in the monkey brain. *Nat. Neurosci.* 11, 367–374.
- Perrodin, C., Kayser, C., Logothetis, N.K., and Petkov, C.I. (2011). Voice cells in the primate temporal lobe. *Curr. Biol.* 21, 1408–1415.
- Poremba, A., Malloy, M., Saunders, R.C., Carson, R.E., Herscovitch, P., and Mishkin, M. (2004). Species-specific calls evoke asymmetric activity in the monkey's temporal poles. *Nature* 427, 448–451.
- Gil-da-Costa, R., Martin, A., Lopes, M.A., Muñoz, M., Fritz, J.B., and Braun, A.R. (2006). Species-specific calls activate homologs of Broca's and Wernicke's areas in the macaque. *Nat. Neurosci.* 9, 1064–1070.
- Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M.L., Scherer, K.R., and Vuilleumier, P. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nat. Neurosci.* 8, 145–146.
- Ethofer, T., Van De Ville, D., Scherer, K.R., and Vuilleumier, P. (2009). Decoding of emotional information in voice-sensitive cortices. *Curr. Biol.* 19, 1028–1033.
- Overath, T., Kumar, S., Stewart, L., von Kriegstein, K., Cusack, R., Rees, A., and Griffiths, T.D. (2010). Cortical mechanisms for the segregation and representation of acoustic textures. *J. Neurosci.* 30, 2070–2076.
- Belin, P., Fecteau, S., Charest, I., Nicastro, N., Hauser, M.D., and Armony, J.L. (2008). Human cerebral response to animal affective vocalizations. *Proc. Biol. Sci.* 275, 473–481.
- Morton, E.S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *Am. Nat.* 111, 855–869.
- Faragó, T., Andics, A., Devecseri, V., Kis, A., Gácsi, M., and Miklósi, A. (2014). Humans rely on the same rules to assess emotional valence and intensity in conspecific and dog vocalizations. *Biol. Lett.* 10, 20130926.
- August, P.V., and Anderson, J.G.T. (1987). Mammal sounds and motivation-structural rules: a test of the hypothesis. *J. Mammal.* 68, 1–9.
- Lemasson, A., Remeuf, K., Rossard, A., and Zimmermann, E. (2012). Cross-taxa similarities in affect-induced changes of vocal behavior and voice in arboreal monkeys. *PLoS ONE* 7, e45106.
- Fichtel, C., Hammerschmidt, K., and Jürgens, U. (2001). On the vocal expression of emotion. A multi-parametric analysis of different states of aversion in the squirrel monkey. *Behaviour* 138, 97–116.
- Rendall, D. (2003). Acoustic correlates of caller identity and affect intensity in the vowel-like grunt vocalizations of baboons. *J. Acoust. Soc. Am.* 113, 3390–3402.
- Scheumann, M., Roser, A.-E., Konerding, W., Bleich, E., Hedrich, H.-J., and Zimmermann, E. (2012). Vocal correlates of sender-identity and arousal in the isolation calls of domestic kitten (*Felis silvestris catus*). *Front. Zool.* 9, 36.
- Graham, M.A., and Noonan, M. (2010). Call types and acoustic features associated with aggressive chase in the killer whale (*Orcinus orca*). *Aquat. Mamm.* 36, 9–18.
- Briefer, E.F. (2012). Vocal expression of emotions in mammals: mechanisms of production and evidence. *J. Zool. (Lond.)* 288, 1–20.
- Taylor, A.M., and Reby, D. (2010). The contribution of source-filter theory to mammal vocal communication research. *J. Zool. (Lond.)* 280, 221–236.
- Owings, D.H., and Morton, E.S. (1998). *Animal Vocal Communication: A New Approach* (Cambridge: Cambridge University Press).
- Tóth, L., Gácsi, M., Miklósi, Á., Bogner, P., and Repa, I. (2009). Awake dog brain magnetic resonance imaging. *J. Vet. Behav. Clin. Appl. Res.* 4, 50.
- Berns, G.S., Brooks, A.M., and Spivak, M. (2012). Functional MRI in awake unrestrained dogs. *PLoS ONE* 7, e38027.
- Heffner, H. (1978). Effect of auditory cortex ablation on localization and discrimination of brief sounds. *J. Neurophysiol.* 41, 963–976.
- Stepień, I., Stepień, L., and Lubińska, E. (1990). Function of dog's auditory cortex in tests involving auditory location cues and directional instrumental response. *Acta Neurobiol. Exp. (Warsz.)* 50, 1–12.
- Szwejkowska, G., and Sychowa, B. (1971). The effects of lesions of auditory cortex on discrimination of sound localization in dog. *Acta Neurobiol. Exp. (Warsz.)* 31, 237–250.
- von Kriegstein, K., Smith, D.R.R., Patterson, R.D., Ives, D.T., and Griffiths, T.D. (2007). Neural representation of auditory size in the human voice and in sounds from other resonant sources. *Curr. Biol.* 17, 1123–1128.
- Joly, O., Pallier, C., Ramus, F., Pressnitzer, D., Vanduffel, W., and Orban, G.A. (2012). Processing of vocalizations in humans and monkeys: a comparative fMRI study. *Neuroimage* 62, 1376–1389.
- Joly, O., Ramus, F., Pressnitzer, D., Vanduffel, W., and Orban, G.A. (2012). Interhemispheric differences in auditory processing revealed by fMRI in awake rhesus monkeys. *Cereb. Cortex* 22, 838–853.
- Belin, P., and Grosbras, M.-H. (2010). Before speech: cerebral voice processing in infants. *Neuron* 65, 733–735.
- Kriegstein, K.V., and Giraud, A.-L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *Neuroimage* 22, 948–955.
- von Kriegstein, K., Eger, E., Kleinschmidt, A., and Giraud, A.L. (2003). Modulation of neural responses to speech by directing attention to voices or verbal content. *Brain Res. Cogn. Brain Res.* 17, 48–55.
- Fitch, W.T., Huber, L., and Bugnyar, T. (2010). Social cognition and the evolution of language: constructing cognitive phylogenies. *Neuron* 65, 795–814.
- Fecteau, S., Armony, J.L., Joanette, Y., and Belin, P. (2004). Is voice processing species-specific in human auditory cortex? An fMRI study. *Neuroimage* 23, 840–848.
- Fecteau, S., Armony, J.L., Joanette, Y., and Belin, P. (2005). Sensitivity to voice in human prefrontal cortex. *J. Neurophysiol.* 94, 2251–2254.
- Andics, A., McQueen, J.M., and Petersson, K.M. (2013). Mean-based neural coding of voices. *Neuroimage* 79, 351–360.
- Latinus, M., Crabbe, F., and Belin, P. (2011). Learning-induced changes in the cerebral processing of voice identity. *Cereb. Cortex* 21, 2820–2828.
- von Kriegstein, K., Patterson, R.D., and Griffiths, T.D. (2008). Task-dependent modulation of medial geniculate body is behaviorally relevant for speech recognition. *Curr. Biol.* 18, 1855–1859.
- Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., and Ethofer, T. (2008). Cerebral processing of emotional prosody— influence of acoustic parameters and arousal. *Neuroimage* 39, 885–893.
- Latinus, M., and Belin, P. (2011). Human voice perception. *Curr. Biol.* 21, R143–R145.
- Custance, D., and Mayer, J. (2012). Empathic-like responding by domestic dogs (*Canis familiaris*) to distress in humans: an exploratory study. *Anim. Cogn.* 15, 851–859.
- Schirmer, A., and Kotz, S.A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn. Sci.* 10, 24–30.
- Siniscalchi, M., Quaranta, A., and Rogers, L.J. (2008). Hemispheric specialization in dogs for processing different acoustic stimuli. *PLoS ONE* 3, e3349.